

AMENDMENTS TO THE CLAIMS

Please amend the claims as follows:

1 1. (Cancelled)

2 2. (Currently amended) ~~The method of claim 1, further comprising, prior~~

3 ~~to d) A computer-implemented method for randomly walking through a hyper-~~

4 ~~text-linked document set comprising a plurality of documents, wherein at least a~~

5 ~~subset of the documents contain a plurality of links to other documents, each~~

6 ~~document being associated with a host, the method comprising:~~

7 a) ~~selecting a host;~~

8 b) ~~selecting at random a document associated with the host;~~

9 c) ~~retrieving the selected document;~~

10 e.1 d) ~~responsive to occurrence of a random event:~~

11 e.1.1 d.1) ~~selecting at random a host from among the pre-~~

12 viously selected hosts; and

13 d.2) ~~selecting at random a document associated with the~~

14 host; and

15 d.3) ~~retrieving the selected document;~~

16 e.1.2) ~~repeating b) through f);~~

e) ~~responsive to non-occurrence of the random event:~~

17 e.1) selecting at random a link in the retrieved document;

18 and

19 e.2) retrieving a document referenced by the selected link;

20 and

21 f) repeating d) and e) until a predetermined condition is met.

22 and wherein f) comprises repeating e.1) through e) until a predetermined

23 condition is met.

1 3. (Currently amended) The method of claim 1, further comprising, prior

2 to d) 2, wherein the random event comprises:

3 e.1) generating a random number;

4 e.2) determining whether the a generated random number falls

5 falling within a predetermined range; and

6 e.3) responsive to the random number falling within the prede-

7 termined range:

8 e.1.1) selecting at random a host from among the previously

9 selected hosts; and

10 e.1.2) repeating b) through f).

1 4. (Currently Amended) The method of claim 1, 2, wherein the document

2 set is the World Wide Web, and wherein each document is a web page.

1 5. (Original) The method of claim 4, wherein each host corresponds to a
2 domain.

1 6. (Currently amended) The method of claim 4, further comprising, con-
2 currently with a) through f), performing a second two-level random walk through
3 the hypertext-linked document set.

1 7. (Currently amended) A computer-implemented method for randomly
2 walking through a hypertext-linked document set comprising a plurality of
3 documents, wherein at least a subset of the documents contain a plurality of links
4 to other documents, each document being associated with a host, the method
5 comprising:

- 6 a) initializing a host set;
- 7 b) initializing a document set for each host in the host set;
- 8 c) selecting at random a host from the host set;
- 9 d) selecting at random a document from the document set of the
10 selected host;
- 11 e) ~~adding the selected host to the host set;~~
- 12 f) ~~adding the selected document to the document set of the se-
13 lected host;~~
- 14 g) ~~e)~~ responsive to the selected document containing at least one
15 link:

16 g.1 e.1) selecting at random a link from the selected doc-
17 ument;
18 g.2 e.2) selecting a document corresponding to the se-
19 lected link;
20 g.3 e.3) selecting a host corresponding to the selected
21 document;
22 e.4) adding the selected host to the host set;
23 e.5) adding the selected document to the document set of
24 the selected host; and
25 g.4 e.6) repeating e.1) through h.5) until a first prede-
26 termined condition is met; and
27 h. f) responsive to the selected document not containing at least
28 one link, repeating c) through h. e) until a second predeter-
29 mined condition is met.

1 8. (Currently amended) The method of claim 7, wherein:

2 e.4) is performed responsive to the selected host not being in the host

3 set; and

4 f e.5) is performed responsive to the selected document not being in

5 the document set of the selected host.

1 9. (Currently amended) ~~The method of claim 7, wherein g) further com-~~
2 ~~prises, prior to g.1) A computer-implemented method for randomly walking~~

3 through a hypertext-linked document set comprising a plurality of documents,
4 wherein at least a subset of the documents contain a plurality of links to other
5 documents, each document being associated with a host, the method comprising:

6 a) initializing a host set;
7 b) initializing a document set for each host in the host set;
8 c) selecting at random a host from the host set;
9 d) selecting at random a document from the document set of the
10 selected host;
11 e) responsive to non-occurrence of a random event, and further
12 responsive to the selected document containing at least one
13 link:
14 e.1) selecting at random a link from the selected document;
15 e.2) selecting a document corresponding to the selected
16 link;
17 e.3) selecting a host corresponding to the selected doc-
18 ument;
19 e.4) adding the selected host to the host set;
20 e.5) adding the selected document to the document set of
21 the selected host; and
22 e.6) repeating e.1) through e.5) until a first predetermined
23 condition is met; and

24 f) repeating c) through e) until a second predetermined con-
25 dition is met

26 g.0) ~~responsive to a random event, repeating c) through h) until a~~
27 predetermined condition is met;

28 and wherein g.1) through g.4) are performed responsive to non-occurrence
29 of the random event of g.0).

1 10. (Currently amended) The method of claim 7, further comprising, prior

2 to g.1): 9, wherein the random event comprises:

3 g.0.1) ~~generating a random number;~~

4 g.0.2) ~~determining whether the a generated random number falls~~
5 falling within a predetermined range; and

6 g.0.3) ~~responsive to the random number falling within the prede-~~
7 termined range, repeating c) through h) until a predetermined
8 condition is met;

9 and wherein g.1) through g.4) are performed responsive to the random
10 number not falling within a predetermined range.

1 11. (Original) The method of claim 7, wherein the hypertext-linked docu-
2 ment set is the World Wide Web, and wherein each document is a web page.

1 12. (Original) The method of claim 11, wherein each host corresponds to a
2 domain.

1 13. (Original) A computer-implemented method for measuring relative
2 quality of a search engine index, comprising:

- 3 a) performing a two-level random walk among documents
4 within a document set;
- 5 b) for each document encountered in the random walk, deter-
6 mining whether the document is indexed by the search engine
7 index; and
- 8 c) aggregating the results of b).

1 14. (Original) The method of claim 13, wherein at least a subset of the
2 documents contain a plurality of links to other documents, each document being
3 associated with a host, and wherein a) comprises:

- 4 a.1) selecting a host;
- 5 a.2) selecting at random a document associated with the host;
- 6 a.3) retrieving the selected document;
- 7 a.4) selecting at random a link in the retrieved document;
- 8 a.5) retrieving a document referenced by the selected link; and
- 9 a.6) repeating a.4) and a.5) until a predetermined condition is met.

1 15. (Currently amended) ~~The method of claim 14, further comprising,~~
2 ~~prior to a.4): A computer-implemented method for measuring relative quality of a~~
3 ~~search engine index, comprising:~~

4 a) performing a two-level random walk among documents

5 within a document set, by:

6 a.1) selecting a host;

7 a.2) selecting at random a document associated with the

8 host;

9 a.3) retrieving the selected document;

10 a.3.1) responsive to occurrence of a random event:

11 a.3.1.1) selecting at random a host from among
12 the previously selected hosts; and

13 a.3.1.2) selecting at random a document associ-
14 ated with the host; and

15 a.3.1.3) retrieving the selected document;

16 a.3.1.2) repeating a.2) through a.6);

17 a.3.2) responsive to non-occurrence of the random event:

18 a.4) selecting at random a link in the retrieved
19 document; and

20 a.5) retrieving a document referenced by the selected
21 link; and

22 a.6) repeating a.3.1) through a.5) until a predetermined
23 condition is met;
24 b) for each document encountered in the random walk, deter-
25 mining whether the document is indexed by the search engine
26 index; and
27 c) aggregating the results of b).

1 16. (Original) The method of claim 13, wherein at least a subset of the
2 documents contain a plurality of links to other documents, each document being
3 associated with a host, and wherein a) comprises:

4 a.1) initializing a host set;
5 a.2) initializing a document set for each host in the host set;
6 a.3) selecting at random a host from the host set;
7 a.4) selecting at random a document from the document set of the
8 selected host;
9 a.5) adding the selected host to the host set;
10 a.6) adding the selected document to the document set of the se-
11 lected host;
12 a.7) responsive to the selected document containing at least one
13 link:
14 a.7.1) selecting at random a link from the selected document;

15 a.7.2) selecting a document corresponding to the selected
16 link;
17 a.7.3) selecting a host corresponding to the selected doc-
18 ument;
19 a.7.4) repeating a.5) through a.8) until a predetermined con-
20 dition is met; and
21 a.8) responsive to the selected document not containing at least
22 one link, repeating a.3) through a.8) until a predetermined
23 condition is met.

AM

1 17. (Original) The method of claim 16, wherein:
2 a.5) is performed responsive to the selected host not being in the host
3 set; and
4 a.6) is performed responsive to the selected document not being in
5 the document set of the selected host.

1 18. (Original) The method of claim 13, wherein each document contains a
2 plurality of words, and wherein b) comprises, for each document encountered in
3 the random walk:
4 b.1) selecting at least one word from the document;
5 b.2) performing a query on the search engine index based on the
6 selected at least one word, to obtain search results; and

7 b.3) determining whether the document is included in the ob-
8 tained search results.

1 19. (Original) The method of claim 18, wherein b.1) comprises selecting at
2 least one word based on rarity.

1 20. (Currently amended) A computer-implemented method for measuring
2 relative quality of a target document in a document set, comprising:
3 a) performing a two-level random walk among documents
4 within a document set; and
5 b) determining a quality metric responsive to the number of
6 times the target document is encountered in the random walk.

1 21. (Currently amended) A computer-implemented method for measuring
2 relative quality of a target document in a document set comprising a plurality of
3 documents, wherein at least a subset of the documents contain a plurality of links
4 to other documents, the method comprising:
5 a) performing a two-level random walk among documents
6 within a document set; and
7 b) determining a quality metric responsive to the number of
8 documents encountered during the two-level random walk
9 that link to the target document.

1 22. (Currently amended) The method of claim 21, wherein b) comprises
2 determining a quality metric responsive to the number of documents that link to
3 the target document, and responsive to the quality metric of the linking docu-
4 ments.

1 23. (Currently amended) The method of claim 21, wherein b) comprises
2 determining a value for:

3
$$R(p) = d / T + (1 - d) \sum_{i=1}^k R(p_i) / C(p_i)$$

4 where:

5 R(p) is the PageRank of target document p;

6 R(p_i) is the PageRank of document p_i;

7 T is the total number of documents in the document set;

8 d is a damping factor such that 0 < d < 1;

9 documents p₁, ..., p_k each contain at least one link to target document p;

10 and

11 C(p_i) is the number of links out of document p_i.

1 24. (Currently amended) The method of claim 21, A computer-
2 implemented method for measuring relative quality of a target document in a
3 document set comprising a plurality of documents, wherein at least a subset of the
4 documents contain a plurality of links to other documents, wherein each docu-
5 ment is associated with a host, and wherein a) comprises the method comprising:

6 a) performing a two-level random walk among documents
7 within a document set, by:
8 a.1) selecting a host;
9 a.2) selecting at random a document associated with the
10 host;
11 a.3) retrieving the selected document;
12 a.4) responsive to occurrence of a random event:
13 a.4.1) selecting at random a host from among the pre-
14 viously selected hosts; and
15 a.4.2) selecting at random a document associated with
16 the host; and
17 a.4.3) retrieving the selected document;
18 a.4.2) ~~repeating a.2) through a.7);~~
19 a.5) responsive to non-occurrence of the random event:
20 a.5 a.5.1) selecting at random a link in the retrieved
21 document; and
22 a.6 a.5.2) retrieving a document referenced by the
23 selected link; and
24 a.7 a.6) repeating a.4) to a.6 a.5) until a predetermined
25 condition is met; and

26 b) determining a quality metric responsive to the number of
27 documents encountered during the two-level random walk
28 that link to the target document.

1 25. (Currently amended) The method of claim 21, A computer-
2 implemented method for measuring relative quality of a target document in a
3 document set comprising a plurality of documents, wherein at least a subset of the
4 documents contain a plurality of links to other documents, wherein each docu-
5 ment is associated with a host, and wherein a) comprises: the method comprising:
6 a) performing a two-level random walk among documents
7 within a document set, by:
8 a.1) initializing a host set;
9 a.2) initializing a document set for each host in the host set;
10 a.3) selecting at random a host from the host set;
11 a.4) responsive to occurrence of a random event:
12 a.4.1) selecting at random a host from among the pre-
13 viously selected hosts; and
14 a.4.2) repeating a.2) through a.7).
15 a.5) responsive to non-occurrence of the random event:
16 a.5 a.5.1) selecting at random a document from the
17 document set of the selected host; and

18 a.6) ~~adding the selected host to the host set;~~

19 a.7) ~~adding the selected document to the document~~

20 ~~set of the selected host;~~

21 a.8 a.5.2) responsive to the selected document

22 containing at least one link:

23 a.8.1 a.5.2.1) selecting at random a link from the

24 selected document;

25 a.8.2 a.5.2.2) selecting a document correspond-

26 ing to the selected link;

27 a.8.3 a.5.2.3) selecting a host corresponding to

28 the selected document; and

29 a.5.2.4) ~~adding the selected host to the host~~

30 ~~set;~~

31 a.5.2.5) ~~adding the selected document to~~

32 ~~the document set of the selected host;~~

33 a.8.4 a.5.2.6) repeating a.6 a.5.2.1) through a.9

34 a.5.2.5) until a first predetermined condi-

35 tion is met; and

36 a.9 a.6) ~~responsive to the selected document not con-~~

37 ~~taining at least one link,~~ repeating a.3) through a.9 a.5)

38 until a second predetermined condition is met; and

b) determining a quality metric responsive to the number of documents encountered during the two-level random walk that link to the target document.

1 26. (Currently amended) The method of claim 21, further comprising:

2 c) determining a quality metric for at least one additional target

3 document; and

4 d) ranking the quality metric of the first target document with

5 respect to the quality metrics of the additional target docu-

6 ments.

1 27. (Currently amended) A computer-implemented method for randomly
2 walking through a hypertext-linked document set comprising a plurality of
3 documents, wherein at least a subset of the documents contain a plurality of links
4 to other documents, each document being associated with a host, the method
5 comprising:

d.2) repeating b) through e) until a predetermined condition is met; and

e) responsive to non-occurrence of the random event not occurring:

e.1) selecting at random a link in the retrieved document;

e.2) retrieving a document referenced by the selected link;
and

e.3) repeating d) and e) until a predetermined condition is met.

28. (Currently amended) A computer-implemented method for measuring the quality of a target document in a document set comprising a plurality of documents, wherein at least a subset of the documents contain a plurality of links to other documents, the method comprising:

a) performing a two-level random walk among documents
within a document set, the two-level random walk comprises
ing by:

- a.1) initializing a host set;
- a.2) initializing a document set for each host in the host set;
- a.3) selecting at random a host from the host set;
- a.4) responsive to occurrence of a random event:

12 a.4.1) selecting at random a host from among the pre-
13 viously selected hosts; and

14 a.4.2) ~~repeating a.2) through a.7).~~

15 a.5) responsive to non-occurrence of the random event:

16 a.5 a.5.1) selecting at random a document from the
17 document set of the selected host; and

18 a.6) ~~adding the selected host to the host set;~~

19 a.7) ~~adding the selected document to the document
20 set of the selected host;~~

21 a.8 a.5.2) responsive to the selected document
22 containing at least one link:

23 a.8.1 a.5.2.1) selecting at random a link from the
24 selected document;

25 a.8.2 a.5.2.2) selecting a document correspond-
26 ing to the selected link;

27 a.8.3 a.5.2.3) selecting a host corresponding to
28 the selected document; and

29 a.5.2.4) adding the selected host to the host
30 set;

31 a.5.2.5) adding the selected document to
32 the document set of the selected host;

33 a.8.4 a.5.2.6) repeating a.6 a.5.2.1) through a.9

34 a.5.2.5) until a first predetermined condi-

35 tion is met; and

36 a.9 a.6) responsive to the selected document not con-

37 taining at least one link, repeating a.3) through a.9 a.5)

38 until a second predetermined condition is met; and

39 b) determining a quality metric responsive to the number of

40 documents encountered during the two-level random walk

41 that link to the target document;

42 c) determining a quality metric for at least one additional target

43 document; and

44 d) ranking the quality metric of the first document with respect

45 to the quality metrics of the additional target documents.

AM

1 29. (Cancelled)

1 30. (Currently amended) The computer program product of claim 29, fur-

2 ther comprising computer readable program code devices configured to cause a

3 computer to, prior to selecting at random a link in the retrieved document A

4 computer program product comprising a computer-usuable medium having com-

5 puter-readable code embodied therein for randomly walking through a hypertext-

6 linked document set comprising a plurality of documents, wherein at least a sub-

7 set of the documents contain a plurality of links to other documents, each docu-
8 ment being associated with a host, the computer program product comprising:

9 a) computer-readable program code devices configured to cause
10 a computer to select a host;

11 b) computer-readable program code devices configured to cause
12 a computer to select at random a document associated with
13 the host;

14 c) computer-readable program code devices configured to cause
15 a computer to retrieve the selected document;

16 e.1 d) computer-readable program code devices configured to cause
17 a computer to, responsive to occurrence of a random event:

18 d.1) select at random a host from among the previously se-
19 lected hosts; and

20 d.2) select at random a document associated with the host;
21 and

22 d.3) retrieve the selected document;
23 ~~repeat the operations of b) through f);~~

24 e) computer-readable program code devices configured to cause
25 a computer to, responsive to non-occurrence of the random
26 event:

27 e.1) select at random a link in the retrieved document; and

28 e.2) retrieve a document referenced by the selected link;
29 and
30 f) computer-readable program code devices configured to cause
31 a computer to repeat the operations of d) and e) until a pre-
32 determined condition is met.

33 and wherein the computer readable program code devices configured to
34 cause a computer to repeat the operations of d) and e) until a predetermined con-
35 dition is met comprise computer readable program code devices configured to
36 cause a computer to repeat the operations of c.1) through e) until a predetermined
37 condition is met.

1 31. (Currently amended) The computer program product of claim 29 30,
2 wherein the random event comprises further comprising:
3 computer readable program code devices configured to cause a com-
4 puter to generate a random number;
5 computer readable program code devices configured to cause a com-
6 puter to determine whether the a generated random number
7 falls falling within a predetermined range; and
8 computer readable program code devices configured to cause a com-
9 puter to, responsive to the random number falling within the
10 predetermined range;

11 select at random a host from among the previously selected
12 hosts; and
13 repeat the operations of b) through f).

1 32. (Currently amended) The computer program product of claim 29 30,
2 wherein the document set is the World Wide Web, and wherein each document is
3 a web page.

1 33. (Original) The computer program product of claim 32, wherein each
2 host corresponds to a domain.

1 34. (Original) The computer program product of claim 29 30, further com-
2 prising computer-readable program code devices configured to cause a computer
3 to, concurrently with the operations of a) through f), perform a second two-level
4 random walk through the hypertext-linked document set.

1 35. (Currently amended) A computer program product comprising a com-
2 puter-readable medium having computer-readable code embodied therein for ran-
3 domly walking through a hypertext-linked document set comprising a plurality of
4 documents, wherein at least a subset of the documents contain a plurality of links
5 to other documents, each document being associated with a host, the computer
6 program product comprising:

7 a) computer-readable program code devices configured to cause
8 a computer to initialize a host set;

9 b) computer-readable program code devices configured to cause
10 a computer to initialize a document set for each host in the
11 host set;

12 c) computer-readable program code devices configured to cause
13 a computer to select at random a host from the host set;

14 d) computer-readable program code devices configured to cause
15 a computer to select at random a document from the docu-
16 ment set of the selected host;

17 e) ~~computer-readable program code devices configured to cause~~
18 ~~a computer to add the selected host to the host set;~~

19 f) ~~computer-readable program code devices configured to cause~~
20 ~~a computer to add the selected document to the document set~~
21 ~~of the selected host;~~

22 g) computer-readable program code devices configured to cause
23 a computer to, responsive to the selected document contain-
24 ing at least one link:
25 g.1 e.1) select at random a link from the selected docu-
26 ment;
27 g.2 e.2) select a document corresponding to the selected
28 link;

29 g.3 e.3) select a host corresponding to the selected docu-
30 ment; and

31 e.4) add the selected host to the host set;

32 e.5) add the selected document to the document set of the
33 selected host; and

34 g.4 e.6) repeat the operations of e e.1) through h e.5) un-
35 til a first predetermined condition is met; and

36 h) computer-readable program code devices configured to cause
37 a computer to, ~~responsive to the selected document not con-~~
38 ~~taining at least one link,~~ repeat the operations of c) through h
39 e) until a second predetermined condition is met.

1 36. (Original) The computer program product of claim 35, wherein:

2 the computer-readable program code devices configured to cause a

3 computer to add the selected host to the host set operate re-

4 sponsive to the selected host not being in the host set; and

5 the computer-readable program code devices configured to cause a

6 computer to add the selected document to the document set

7 of the selected host operate responsive to the selected docu-

8 ment not being in the document set of the selected host.

1 37. (Currently amended) The computer program product of claim 35,
2 wherein computer readable program code devices g) further comprise computer-

3 ~~readable program code devices configured to cause a computer to, prior to g.1) A~~

4 ~~computer program product comprising a computer-usable medium having com-~~

5 ~~puter-readable code embodied therein for randomly walking through a hypertext-~~

6 ~~linked document set comprising a plurality of documents, wherein at least a sub-~~

7 ~~set of the documents contain a plurality of links to other documents, each docu-~~

8 ~~ment being associated with a host, the computer program product comprising:~~

9 a) ~~computer-readable program code devices configured to cause~~

10 ~~a computer to initialize a host set;~~

11 b) ~~computer-readable program code devices configured to cause~~

12 ~~a computer to initialize a document set for each host in the~~

13 ~~host set;~~

14 c) ~~computer-readable program code devices configured to cause~~

15 ~~a computer to select at random a host from the host set;~~

16 d) ~~computer-readable program code devices configured to cause~~

17 ~~a computer to select at random a document from the docu-~~

18 ~~ment set of the selected host;~~

19 e) ~~computer-readable program code devices configured to cause~~

20 ~~a computer to, responsive to non-occurrence of a random~~

21 ~~event, and further responsive to the selected document con-~~

22 ~~taining at least one link:~~

23 e.1) ~~select at random a link from the selected document;~~

24 e.2) ~~select a document corresponding to the selected link;~~

25 e.3) select a host corresponding to the selected document;

26 and

27 e.4) add the selected host to the host set;

28 e.5) add the selected document to the document set of the
29 selected host; and

30 e.6) repeat the operations of e.1) through e.5) until a first
31 predetermined condition is met; and

32 f) computer-readable program code devices configured to cause
33 a computer to repeat the operations of c) through e) until a
34 second predetermined condition is met.

35 g.0) responsive to a random event, repeat the operations of c)
36 through h) until a predetermined condition is met;

37 and wherein computer readable program code devices g) are configured to
38 cause a computer to perform g.1) through g.4) responsive to non occurrence of the
39 random event of g.0).

1 38. (Currently amended) The computer program product of claim 35,
2 wherein computer readable program code devices g) further comprise computer-
3 readable program code devices configured to cause a computer to, prior to g.1) 37,
4 wherein the random event comprises:

5 g.0.1) generate a random number;

6 g.0.2) determine whether the a generated random number falls fal-
7 ling within a predetermined range; and
8 g.0.3) responsive to the random number falling within the prede-
9 termined range, repeat the operations of c) through h) until a
10 predetermined condition is met;

11 and wherein computer readable program code devices g) are configured to
12 cause a computer to perform g.1) through g.4) responsive to the random number
13 not falling within a predetermined range.

1 39. (Original) The computer program product of claim 35, wherein the hy-
2 pertext-linked document set is the World Wide Web, and wherein each document
3 is a web page.

1 40. (Original) The computer program product of claim 39, wherein each
2 host corresponds to a domain.

1 41. (Original) A computer program product comprising a computer-usa-
2 ble medium having computer-readable code embodied therein for measuring relative
3 quality of a search engine index, the computer program product comprising:
4 a) computer-readable program code devices configured to cause
5 a computer to perform a two-level random walk among
6 documents within a document set;

1 42. (Original) The computer program product of claim 41, wherein at least
2 a subset of the documents contain a plurality of links to other documents, each
3 document being associated with a host, and wherein the computer-readable pro-
4 gram code devices configured to cause a computer to perform a two-level random
5 walk comprise:

16 a.5) computer-readable program code devices configured to cause
17 a computer to retrieve a document referenced by the selected
18 link; and
19 a.6) computer-readable program code devices configured to cause
20 a computer to repeat the operations of a.4) and a.5) until a
21 predetermined condition is met.

1 43. (Currently amended) ~~The computer program product of claim 42, fur-~~
2 ~~ther comprising computer readable program code devices configured to cause a~~
3 ~~computer to, prior to selecting at random a link in the retrieved document A~~
4 ~~computer program product comprising a computer-readable medium having com-~~
5 ~~puter-readable code embodied therein for measuring relative quality of a search~~
6 ~~engine index, the computer program product comprising:~~

7 a) ~~computer-readable program code devices configured to cause~~
8 ~~a computer to perform a two-level random walk among~~
9 ~~documents within a document set, wherein at least a subset of~~
10 ~~the documents contain a plurality of links to other documents,~~
11 ~~each document being associated with a host, the two-level~~
12 ~~random walk comprising:~~
13 a.1) ~~selecting a host;~~
14 a.2) ~~selecting at random a document associated with the~~
15 ~~host;~~

16 a.3) retrieving the selected document;

17 a.3.1) responsive to occurrence of a random event:

18 a.3.1.1) selecting at random a host from among

19 the previously selected hosts; and

20 a.3.1.2) selecting at random a document associ-

21 ated with the host; and

22 a.3.1.3) retrieving the selected document;

23 repeat the operations of a.2) through a.6)

24 a.3.2) responsive to non-occurrence of the random event:

25 a.4) selecting at random a link in the retrieved

26 document; and

27 a.5) retrieving a document referenced by the selected

28 link; and

29 a.6) repeating a.3.1) through a.5) until a predetermined

30 condition is met;

31 b) computer-readable program code devices configured to cause

32 a computer to, for each document encountered in the random

33 walk, determine whether the document is indexed by the

34 search engine index; and

35 c) computer-readable program code devices configured to cause

36 a computer to aggregate the results of the operations of b).

1 44. (Original) The computer program product of claim 41, wherein at least
2 a subset of the documents contain a plurality of links to other documents, each
3 document being associated with a host, and wherein the computer-readable pro-
4 gram code devices configured to cause a computer to perform a two-level random
5 walk comprise:

6 a.1) computer-readable program code devices configured to cause
7 a computer to initialize a host set;

8 a.2) computer-readable program code devices configured to cause
9 a computer to initialize a document set for each host in the
10 host set;

11 a.3) computer-readable program code devices configured to cause
12 a computer to select at random a host from the host set;

13 a.4) computer-readable program code devices configured to cause
14 a computer to select at random a document from the docu-
15 ment set of the selected host;

16 a.5) computer-readable program code devices configured to cause
17 a computer to add the selected host to the host set;

18 a.6) computer-readable program code devices configured to cause
19 a computer to add the selected document to the document set
20 of the selected host;

21 a.7) computer-readable program code devices configured to cause
22 a computer to, responsive to the selected document contain-
23 ing at least one link:
24 a.7.1) select at random a link from the selected document;
25 a.7.2) select a document corresponding to the selected link;
26 a.7.3) select a host corresponding to the selected document;
27 a.7.4) repeat the operations of a.5) through a.8) until a prede-
28 termined condition is met; and
29 a.8) computer-readable program code devices configured to cause
30 a computer to, responsive to the selected document not con-
31 taining at least one link, repeat the operations of a.3) through
32 a.8) until a predetermined condition is met.

1 45. (Original) The computer program product of claim 44, wherein:
2 the computer-readable program code devices configured to cause a
3 computer to add the selected host to the host set are config-
4 ured to cause a computer to add the selected host responsive
5 to the selected host not being in the host set; and
6 the computer-readable program code devices configured to cause a
7 computer to add the selected document to the document set
8 of the selected host are configured to cause a computer to add

the selected document responsive to the selected document
not being in the document set of the selected host.

1 46. (Currently amended) The computer program product of claim 41,
2 wherein each document contains a plurality of words, and wherein the computer-
3 readable program code devices configured to cause a computer to, determine
4 whether the document is indexed by the search engine index comprise computer-
5 readable program code devices configured to, for each document encountered in
6 the random walk:

- b.1) select at least one word from the document;
- b.2) perform a query on the search engine index based on the selected at least one word, to obtain search results; and
- b.3) determine whether the document is included in the obtained search results.

1 47. (Original) The computer program product of claim 46, wherein the
2 computer-readable program code devices configured to select at least one word
3 from the document comprise computer-readable program code devices configured
4 to select at least one word based on rarity.

1 48. (Currently amended) A computer program product comprising a com-
2 puter-usable medium having computer-readable code embodied therein for mea-

3 suring relative quality of a target document in a document set, the computer pro-
4 gram product comprising:
5 computer-readable program code devices configured to cause a com-
6 puter to perform a two-level random walk among documents
7 within a document set; and
8 computer-readable program code devices configured to cause a com-
9 puter to determine a quality metric responsive to the number
10 of times the target document is encountered in the random
11 walk.

1 49. (Currently amended) A computer program product comprising a com-
2 puter-readable medium having computer-readable code embodied therein for
3 measuring relative quality of a target document in a document set comprising a
4 plurality of documents, wherein at least a subset of the documents contain a plu-
5 rality of links to other documents, the computer program product comprising:
6 computer-readable program code devices configured to cause a com-
7 puter to perform a two-level random walk among documents
8 within a document set; and
9 computer-readable program code devices configured to cause a com-
10 puter to determine a quality metric responsive to the number
11 of documents encountered during the two-level random walk
12 that link to the target document.

1 50. (Currently amended) The computer program product of claim 49,
2 wherein the computer-readable program code devices configured to cause a com-
3 puter to determine a quality metric comprise computer-readable program code
4 devices configured to cause a computer to determine a quality metric responsive
5 to the number of documents that link to the target document, and responsive to
6 the quality metric of the linking documents.

1 51. (Currently amended) The computer program product of claim 49,
2 wherein the computer-readable program code devices configured to cause a com-
3 puter to determine a quality metric comprise computer-readable program code
4 devices configured to cause a computer to determine a value for:

5
$$R(p) = d / T + (1 - d) \sum_{i=1}^k R(p_i) / C(p_i)$$

6 where:

7 R(p) is the PageRank of target document p;

8 R(p_i) is the PageRank of document p_i;

9 T is the total number of documents in the document set;

10 d is a damping factor such that 0 < d < 1;

11 documents p₁, ..., p_k each contain at least one link to target document p;

12 and

13 C(p_i) is the number of links out of document p_i.

1 52. (Currently amended) ~~The computer program product of claim 49, A~~
2 ~~computer program product comprising a computer-usable medium having com-~~
3 ~~puter-readable code embodied therein for measuring relative quality of a target~~
4 ~~document in a document set comprising a plurality of documents, wherein at least~~
5 ~~a subset of the documents contain a plurality of links to other documents, and~~
6 ~~wherein each document is associated with a host, and wherein the computer-~~
7 ~~readable program code devices configured to cause a computer to perform a two-~~
8 ~~level random walk comprise the computer program product comprising:~~
9 ~~computer-readable program code devices configured to cause a com-~~
10 ~~puter to perform a two-level random walk among documents~~
11 ~~within a document set, by:~~
12 a.1) ~~computer readable program code devices configured~~
13 ~~to cause a computer to selecting a host;~~
14 a.2) ~~computer readable program code devices configured~~
15 ~~to cause a computer to selecting at random a document~~
16 ~~associated with the host;~~
17 a.3) ~~computer readable program code devices configured~~
18 ~~to cause a computer to retrieve retrieving the selected~~
19 ~~document;~~

20 a.4) ~~computer readable program code devices configured~~
21 ~~to cause a computer to, responsive to occurrence of a~~
22 ~~random event:~~
23 a.4.1) ~~selecting at random a host from among the pre-~~
24 ~~viously selected hosts; and~~
25 a.4.2) ~~selecting at random a document associated with~~
26 ~~the host; and~~
27 a.4.3) ~~retrieving the selected document;~~
28 a.4.2) ~~repeat the operations of a.2) through a.7);~~
29 a.5) responsive to non-occurrence of the random event:
30 a.5 a.6) ~~computer readable program code devices~~
31 ~~configured to cause a computer to selecting at~~
32 ~~random a link in the retrieved document; and~~
33 a.6 a.7) ~~computer readable program code devices~~
34 ~~configured to cause a computer to retrieve re-~~
35 ~~trieving a document referenced by the selected~~
36 ~~link; and~~
37 a.7 a.8) ~~computer readable program code devices con-~~
38 ~~figured to cause a computer to repeating the operations~~
39 ~~of a.4) to a.6 a.7) until a predetermined condition is~~
40 ~~met; and-~~

41 computer-readable program code devices configured to cause a com-
42 puter to determine a quality metric responsive to the number
43 of documents encountered during the two-level random walk
44 that link to the target document.

1 53. (Currently amended) The computer program product of claim 49, A
2 computer program product comprising a computer-usable medium having com-
3 puter-readable code embodied therein for measuring relative quality of a target
4 document in a document set comprising a plurality of documents, wherein at least
5 a subset of the documents contain a plurality of links to other documents, wherein
6 each document is associated with a host, the computer program product compris-
7 ing and wherein and wherein the computer readable program code devices con-
8 figured to cause a computer to perform a two-level random walk comprise:
9 computer-readable program code devices configured to cause a com-
10 puter to perform a two-level random walk among documents
11 within a document set, by:
12 a.1) computer-readable program code devices configured
13 to cause a computer to initialize initializing a host set;
14 a.2) computer-readable program code devices configured
15 to cause a computer to initialize initializing a document
16 set for each host in the host set;

17 a.3) ~~computer readable program code devices configured~~
18 ~~to cause a computer to selecting at random a host from~~
19 ~~the host set;~~

20 a.4) ~~computer readable program code devices configured~~
21 ~~to cause a computer to, responsive to occurrence of a~~
22 ~~random event:~~

23 a.4.1) ~~selecting at random a host from among the pre-~~
24 ~~viously selected hosts; and~~

25 a.4.2) ~~repeat the operations of a.2) through a.7).~~

26 a.5) responsive to non-occurrence of the random event:

27 a.5 a.5.1) ~~computer readable program code devices~~
28 ~~configured to cause a computer to selecting at~~
29 ~~random a document from the document set of~~
30 ~~the selected host;~~

31 a.6 a.5.2) ~~computer readable program code devices~~
32 ~~configured to cause a computer to adding the se-~~
33 ~~lected host to the host set;~~

34 a.7 a.5.3) ~~computer readable program code devices~~
35 ~~configured to cause a computer to adding the se-~~
36 ~~lected document to the document set of the se-~~
37 ~~lected host;~~

38 a.8 a.5.4) ~~computer readable program code devices~~
39 ~~configured to cause a computer to, responsive to~~
40 the selected document containing at least one
41 link:
42 a.8.1 a.5.4.1) ~~selecting at random a link from the~~
43 ~~selected document;~~
44 a.8.2 a.5.4.2) ~~selecting a document correspond-~~
45 ~~ing to the selected link;~~
46 a.8.3 a.5.4.3) ~~selecting a host corresponding to~~
47 ~~the selected document; and~~
48 a.8.4 a.5.4.4) ~~repeating the operations of a.6~~
49 ~~a.5.2) through a.9 a.5.4.3) until a first pre-~~
50 ~~determined condition is met; and~~
51 a.9 a.6) ~~responsive to the selected document not~~
52 ~~containing at least one link, repeating the operations of~~
53 a.3) through a.9 a.5.4.4) until a second predetermined
54 condition is met; and
55 computer-readable program code devices configured to cause a com-
56 puter to determine a quality metric responsive to the number
57 of documents encountered during the two-level random walk
58 that link to the target document.

1 54. (Currently amended) The computer program product of claim 49, fur-
2 ther comprising:

3 c) computer-readable program code devices configured to cause
4 a computer to determine a quality metric for at least one ad-
5 ditional target document; and
6 d) computer-readable program code devices configured to cause
7 a computer to rank the quality metric of the first target docu-
8 ment with respect to the quality metrics of the additional tar-
9 get documents.

1 55. (Currently amended) A computer program product comprising a com-
2 puter-readable medium having computer-readable code embodied therein for ran-
3 domly walking through a hypertext-linked document set comprising a plurality of
4 documents, wherein at least a subset of the documents contain a plurality of links
5 to other documents, each document being associated with a host, the computer
6 program product comprising:

7 a) computer-readable program code devices configured to cause
8 a computer to select a host;
9 b) computer-readable program code devices configured to cause
10 a computer to select at random a document associated with
11 the host;

- c) computer-readable program code devices configured to cause a computer to retrieve the selected document;
- d) computer-readable program code devices configured to cause a computer to, responsive to occurrence of a random event:
 - d.1) select at random a host from among the previously selected hosts; and
 - d.2) repeat the operations of b) through e) until a predetermined condition is met
- e) computer-readable program code devices configured to cause a computer to, responsive to non-occurrence of the random event ~~not occurring~~:
 - e.1) select at random a link in the retrieved document;
 - e.2) retrieve a document referenced by the selected link; and
 - e.3) repeat the operations of d) and e) until a predetermined condition is met.

1 56. (Currently amended) A computer program product comprising a com-
2 puter-usable medium having computer-readable code embodied therein for
3 measuring relative quality of a target document in a document set comprising a
4 plurality of documents, wherein at least a subset of the documents contain a plu-
5 rality of links to other documents, the computer program product comprising:

6 a) computer-readable program code devices configured to cause
7 a computer to perform a two-level random walk among
8 documents within a document set, the computer readable
9 program code devices comprising by:
10 a.1) computer readable program code devices configured to
11 cause a computer to initialize initializing a host set;
12 a.2) computer readable program code devices configured to
13 cause a computer to initialize initializing a document set
14 for each host in the host set;
15 a.3) computer readable program code devices configured to
16 cause a computer to selecting at random a host from the
17 host set;
18 a.4) computer readable program code devices configured to
19 cause a computer to, responsive to occurrence of a random
20 event:
21 a.4.1) selecting at random a host from among the previ-
22 ously selected hosts; and
23 a.4.2) repeat the operations of a.2) through a.7).
24 a.5) responsive to non-occurrence of the random event:
25 a.5 a.5.1) computer readable program code devices con-
26 figured to cause a computer to selecting at random

27 a document from the document set of the selected
28 host;

29 a.6 a.5.2) ~~computer readable program code devices con-~~
30 ~~figured to cause a computer to adding~~ the selected
31 host to the host set;

32 a.7 a.5.3) ~~computer readable program code devices con-~~
33 ~~figured to cause a computer to adding~~ the selected
34 document to the document set of the selected host;

35 a.8 a.5.4) ~~computer readable program code devices con-~~
36 ~~figured to cause a computer to, responsive to the se-~~
37 lected document containing at least one link:

38 a.8.1 a.5.4.1) ~~selecting at random a link from the~~
39 ~~selected document;~~

40 a.8.2 a.5.4.2) ~~selecting a document correspond-~~
41 ~~ing to the selected link;~~

42 a.8.3 a.5.4.3) ~~selecting a host corresponding to~~
43 ~~the selected document;~~

44 a.8.4 a.5.4.4) ~~repeating the operations of a.6~~
45 ~~a.5.2) through a.9 a.5.4.3) until a first pre-~~
46 ~~determined condition is met; and~~

47 a.9) ~~computer readable program code devices configured to~~
48 ~~cause a computer to, responsive to the selected document~~

49 not containing at least one link, repeat the operations of
50 a.3) through a.9 a.5.4.4) until a second predetermined
51 condition is met;

52 b) computer-readable program code devices configured to cause
53 a computer to determine a quality metric responsive to the
54 number of documents encountered during the two-level ran-
55 dom walk that link to the target document;

56 c) computer-readable program code devices configured to cause
57 a computer to determine a quality metric for at least one ad-
58 ditional target document; and

59 d) computer-readable program code devices configured to cause
60 a computer to rank the quality metric of the first document
61 with respect to the quality metrics of the additional target
62 documents.

1 57. (Currently amended) A system for randomly walking through a hyper-
2 text-linked document set comprising a plurality of documents, wherein at least a
3 subset of the documents contain a plurality of links to other documents, each
4 document being associated with a host, the system comprising:
5 a) a host selector;
6 b) a random document selector, coupled to the host selector, for
7 selecting at random a document associated with the host;

- c) a document retriever, coupled to the random document selector, for retrieving the selected document; and

d) a link selector, coupled to the document retriever, for selecting at random a link in the retrieved document;

12 wherein, responsive to occurrence of a random event:

17 the document retriever retrieves the selected document; and

18 wherein, responsive to non-occurrence of the random event:

the link selector selects at random a link in the retrieved document;

20 and

21 the document retriever retrieves a document referenced by the se-
22 lected link; and

23 and wherein the link selector, the random document selector, and the
24 document retriever repeatedly their respective operations selects at random a link
25 and the document retriever repeatedly retrieves a document referenced by the se-
26 lected link, until a predetermined condition is met.

1 58. (Original) A system for measuring relative quality of a search engine
2 index, comprising:
3 a random walker, for performing a two-level random walk among
4 documents within a document set;
5 a determination module, coupled to the random walker, for, for each
6 document encountered in the random walk, determining
7 whether the document is indexed by the search engine index;
8 and
9 a results aggregation module, coupled to the determination module,
10 for aggregating the results of the determination module.



1 59. (Currently amended) A system for measuring relative quality of a tar-
2 get document in a document set, comprising:
3 a random walker, for performing a two-level random walk among
4 documents within a document set; and
5 a determination module, coupled to the random walker, for deter-
6 mining a quality metric responsive to the number of times the
7 target document is encountered in the random walk.
